

RECEIVED

APR 30 2001

TECH CENTER 1600/2900

1645

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/313,434A

DATE: 04/17/2001

TIME: 14:14:02

Input Set : A:\00786-432001.TXT

Output Set: N:\CRF3\04172001\I313434A.raw

ENTERED

4 <110> APPLICANT: Podolsky, Daniel K.
 6 <120> TITLE OF INVENTION: INTESTINAL TREFOIL PROTEINS
 9 <130> FILE REFERENCE: 00786-432001
 11 <140> CURRENT APPLICATION NUMBER: US 09/313,434A
 12 <141> CURRENT FILING DATE: 1999-05-17
 14 <150> PRIOR APPLICATION NUMBER: US 08/631,469
 15 <151> PRIOR FILING DATE: 1996-04-12
 17 <150> PRIOR APPLICATION NUMBER: US 08/191,352
 18 <151> PRIOR FILING DATE: 1994-02-02
 20 <150> PRIOR APPLICATION NUMBER: US 08/037,741
 21 <151> PRIOR FILING DATE: 1993-03-25
 23 <150> PRIOR APPLICATION NUMBER: US 07/837,192
 24 <151> PRIOR FILING DATE: 1992-02-13
 26 <150> PRIOR APPLICATION NUMBER: US 07/655,965
 27 <151> PRIOR FILING DATE: 1991-02-14
 29 <160> NUMBER OF SEQ ID NOS: 21
 31 <170> SOFTWARE: FastSEQ for Windows Version 4.0
 33 <210> SEQ ID NO: 1
 34 <211> LENGTH: 431
 35 <212> TYPE: DNA
 36 <213> ORGANISM: Rattus norvegicus
 38 <220> FEATURE:
 39 <221> NAME/KEY: CDS
 40 <222> LOCATION: (18)...(260)
 42 <400> SEQUENCE: 1
 43 gaagtttgcg tgctgcc / atg gag acc aga gcc ttc tgg ata acc ctg ctg 50
 44 Met Glu Thr Arg Ala Phe Trp Ile Thr Leu Leu
 45 1 5 10
 47 ctg gtc ctg gtt gct ggg tcc tcc tgc aaa gcc cag gaa ttt gtt ggc 98
 48 Leu Val Leu Val Ala Gly Ser Ser Cys Lys Ala Gln Glu Phe Val Gly
 49 15 20 25
 51 cta tct cca agc caa tgt atg gcg cca aca aat gtc agg gtg gac tgt 146
 52 Leu Ser Pro Ser Gln Cys Met Ala Pro Thr Asn Val Arg Val Asp Cys
 53 30 35 40
 55 aac tac ccc act gtc aca tca gag cag tgt aac aac cgt ggt tgc tgt 194
 56 Asn Tyr Pro Thr Val Thr Ser Glu Gln Cys Asn Asn Arg Gly Cys Cys
 57 45 50 55
 59 ttt gac tcc agc atc cca aat gtg ccc tgg tgc ttc aaa cct ctg caa 242
 60 Phe Asp Ser Ser Ile Pro Asn Val Pro Trp Cys Phe Lys Pro Leu Gln
 61 60 65 70 75
 63 gag aca gaa tgt aca ttt / tgaagctgtc caggctccag gaagggagct 290
 64 Glu Thr Glu Cys Thr Phe
 65 80
 67 ccacaccctg gactcttgct gatggtagtg gcccagggta acactcaccc ctgatctgct 350
 68 cctcgcgcgc ggccaatata ggagctggga gtccagaaga ataaagacct tacagtcagc 410
 69 acaaggctgt tctaattgcg g 431
 71 <210> SEQ ID NO: 2

RAW SEQUENCE LISTING

DATE: 04/17/2001

PATENT APPLICATION: US/09/313,434A

TIME: 14:14:02

Input Set : A:\00786-432001.TXT

Output Set: N:\CRF3\04172001\I313434A.raw

```

72 <211> LENGTH: 81
73 <212> TYPE: PRT
74 <213> ORGANISM: Rattus norvegicus
76 <400> SEQUENCE: 2
77 Met Glu Thr Arg Ala Phe Trp Ile Thr Leu Leu Leu Val Leu Val Ala
78 1 5 10 15
79 Gly Ser Ser Cys Lys Ala Gln Glu Phe Val Gly Leu Ser Pro Ser Gln
80 20 25 30
81 Cys Met Ala Pro Thr Asn Val Arg Val Asp Cys Asn Tyr Pro Thr Val
82 35 40 45
83 Thr Ser Glu Gln Cys Asn Asn Arg Gly Cys Cys Phe Asp Ser Ser Ile
84 50 55 60
85 Pro Asn Val Pro Trp Cys Phe Lys Pro Leu Gln Glu Thr Glu Cys Thr
86 65 70 75 80
87 Phe
90 <210> SEQ ID NO: 3
91 <211> LENGTH: 403
92 <212> TYPE: DNA
93 <213> ORGANISM: Homo sapiens
95 <220> FEATURE:
96 <221> NAME/KEY: CDS
97 <222> LOCATION: (2)...(223)
99 <400> SEQUENCE: 3
100 g`atg ctg ggg ctg gtc ctg gcc ttg ctg tcc tcc agc tct gct gag gag 49
101 Met Leu Gly Leu Val Leu Ala Leu Leu Ser Ser Ser Ser Ala Glu Glu
102 1 5 10 15
104 tac gtg ggc ctg tct gca aac cag tgt gcc gtg ccg gcc aag gac agg 97
105 Tyr Val Gly Leu Ser Ala Asn Gln Cys Ala Val Pro Ala Lys Asp Arg
106 20 25 30
108 gtg gac tgc ggc tac ccc cat gtc acc ccc aag gag tgc aac aac cgg 145
109 Val Asp Cys Gly Tyr Pro His Val Thr Pro Lys Glu Cys Asn Asn Arg
110 35 40 45
112 ggc tgc tgc ttt gac tcc agg atc cct gga gtg cct tgg tgt ttc aag 193
113 Gly Cys Cys Phe Asp Ser Arg Ile Pro Gly Val Pro Trp Cys Phe Lys
114 50 55 60
116 ccc ctg act agg aag aca gaa tgc acc ttc`tgaggcacct ccagctgccc 243
117 Pro Leu Thr Arg Lys Thr Glu Cys Thr Phe
118 65 70
120 ctgggatgca ggctgagcac ccttgcccgg ctgtgattgc tgccaggcac tgttcatctc 303
121 agtttttctg tccctttgct cccggcaagc tttctgctga aagttcatat ctggagcctg 363
122 atgtcttaac gaataaaggt cccatgctcc acccgaaaaa 403
124 <210> SEQ ID NO: 4
125 <211> LENGTH: 74
126 <212> TYPE: PRT
127 <213> ORGANISM: Homo sapiens
129 <400> SEQUENCE: 4
130 Met Leu Gly Leu Val Leu Ala Leu Leu Ser Ser Ser Ser Ala Glu Glu
131 1 5 10 15
132 Tyr Val Gly Leu Ser Ala Asn Gln Cys Ala Val Pro Ala Lys Asp Arg

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/313,434A

DATE: 04/17/2001

TIME: 14:14:02

Input Set : A:\00786-432001.TXT

Output Set: N:\CRF3\04172001\I313434A.raw

```

133           20           25           30
134 Val Asp Cys Gly Tyr Pro His Val Thr Pro Lys Glu Cys Asn Asn Arg
135           35           40           45
136 Gly Cys Cys Phe Asp Ser Arg Ile Pro Gly Val Pro Trp Cys Phe Lys
137           50           55           60
138 Pro Leu Thr Arg Lys Thr Glu Cys Thr Phe
139 65           70
141 <210> SEQ ID NO: 5
142 <211> LENGTH: 10
143 <212> TYPE: DNA
144 <213> ORGANISM: Artificial Sequence
146 <220> FEATURE:
147 <223> OTHER INFORMATION: motif
149 <400> SEQUENCE: 5
150 gggcggccgc                                     10
152 <210> SEQ ID NO: 6
153 <211> LENGTH: 21
154 <212> TYPE: DNA
155 <213> ORGANISM: Artificial Sequence
157 <220> FEATURE:
158 <223> OTHER INFORMATION: oligonucleotide for PCR
160 <400> SEQUENCE: 6
161 gtacattctg tctcttgcag a                         21
163 <210> SEQ ID NO: 7
164 <211> LENGTH: 24
165 <212> TYPE: DNA
166 <213> ORGANISM: Artificial Sequence
168 <220> FEATURE:
169 <223> OTHER INFORMATION: oligonucleotide for PCR
171 <400> SEQUENCE: 7
172 taaccctgct gctgctggtc ctgg                       24
174 <210> SEQ ID NO: 8
175 <211> LENGTH: 21
176 <212> TYPE: DNA
177 <213> ORGANISM: Artificial Sequence
179 <220> FEATURE:
180 <223> OTHER INFORMATION: oligonucleotide for PCR
182 <400> SEQUENCE: 8
183 gtttgcgtgc tgccatggag a                         21
185 <210> SEQ ID NO: 9
186 <211> LENGTH: 21
187 <212> TYPE: DNA
188 <213> ORGANISM: Artificial Sequence
190 <220> FEATURE:
191 <223> OTHER INFORMATION: oligonucleotide for PCR
193 <400> SEQUENCE: 9
194 ccgcaattag aacagccttg t                         21
196 <210> SEQ ID NO: 10
197 <211> LENGTH: 25

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/313,434A

DATE: 04/17/2001

TIME: 14:14:02

Input Set : A:\00786-432001.TXT

Output Set: N:\CRF3\04172001\I313434A.raw

```

198 <212> TYPE: DNA
199 <213> ORGANISM: Artificial Sequence
201 <220> FEATURE:
202 <223> OTHER INFORMATION: oligonucleotide for PCR
204 <400> SEQUENCE: 10
205 gcagtgtaac aaccgtgggt gctgc
207 <210> SEQ ID NO: 11
208 <211> LENGTH: 60
209 <212> TYPE: PRT
210 <213> ORGANISM: Homo sapiens
212 <220> FEATURE:
215 <400> SEQUENCE: 11
216 Glu Ala Gln Thr Glu Thr Cys Thr Val Ala Pro Arg Glu Arg Gln Asn
217 1 5 10 15
218 Cys Gly Phe Pro Gly Val Thr Pro Ser Gln Cys Ala Asn Lys Gly Cys
219 20 25 30
220 Cys Phe Asp Asp Thr Val Arg Gly Val Pro Trp Cys Phe Tyr Pro Asn
221 35 40 45
222 Thr Ile Asp Val Pro Pro Glu Glu Glu Cys Glu Phe
223 50 55 60
225 <210> SEQ ID NO: 12
226 <211> LENGTH: 62
227 <212> TYPE: PRT
228 <213> ORGANISM: Homo sapiens
230 <220> FEATURE:
232 <400> SEQUENCE: 12
233 Glu Lys Pro Ala Ala Cys Arg Cys Ser Arg Gln Asp Pro Lys Asn Arg
234 1 5 10 15
235 Val Asn Cys Gly Phe Pro Gly Ile Thr Ser Asp Gln Cys Phe Thr Ser
236 20 25 30
237 Gly Cys Cys Phe Asp Ser Gln Val Pro Gly Val Pro Trp Cys Phe Lys
238 35 40 45
239 Pro Leu Pro Ala Gln Glu Ser Glu Glu Cys Val Met Glu Val
240 50 55 60
242 <210> SEQ ID NO: 13
243 <211> LENGTH: 318
244 <212> TYPE: DNA
245 <213> ORGANISM: Homo sapiens
247 <220> FEATURE:
248 <221> NAME/KEY: CDS
249 <222> LOCATION: (1)...(318)
251 <400> SEQUENCE: 13
252 gag aaa ccc tcc ccc tgc cag tgc tcc agg ctg agc ccc cat aac agg
253 Glu Lys Pro Ser Pro Cys Gln Cys Ser Arg Leu Ser Pro His Asn Arg
254 1 5 10 15
256 acg aac tgc ggc ttc cct gga atc acc agt gac cag tgt ttt gac aat
257 Thr Asn Cys Gly Phe Pro Gly Ile Thr Ser Asp Gln Cys Phe Asp Asn
258 20 25 30
260 gga tgc tgt ttc gac tcc agt gtc act ggg gtc ccc tgg tgt ttc cac

```

RAW SEQUENCE LISTING

DATE: 04/17/2001

PATENT APPLICATION: US/09/313,434A

TIME: 14:14:02

Input Set : A:\00786-432001.TXT

Output Set: N:\CRF3\04172001\I313434A.raw

```

261 Gly Cys Cys Phe Asp Ser Ser Val Thr Gly Val Pro Trp Cys Phe His
262          35          40          45
264 ccc ctc cca aag caa gag tcg gat cag tgc gtc atg gag gtc tca gac      192
265 Pro Leu Pro Lys Gln Glu Ser Asp Gln Cys Val Met Glu Val Ser Asp
266          50          55          60
268 aga aga aac tgt ggc tac ccg ggc atc agc ccc gag gaa tgc gcc tct      240
269 Arg Arg Asn Cys Gly Tyr Pro Gly Ile Ser Pro Glu Glu Cys Ala Ser
270 65          70          75          80
272 cgg aag tgc tgc ttc tcc aac ttc atc ttt gaa gtg ccc tgg tgc ttc      288
273 Arg Lys Cys Cys Phe Ser Asn Phe Ile Phe Glu Val Pro Trp Cys Phe
274          85          90          95
276 ttc ccg aac tct gtg gaa gac tgc cat tac      318
277 Phe Pro Asn Ser Val Glu Asp Cys His Tyr
278          100          105
281 <210> SEQ ID NO: 14
282 <211> LENGTH: 105
283 <212> TYPE: PRT
284 <213> ORGANISM: Homo sapiens
286 <400> SEQUENCE: 14
287 Lys Pro Ser Pro Cys Gln Cys Ser Arg Leu Ser Pro His Asn Arg Thr
288 1          5          10          15
289 Asn Cys Gly Phe Pro Gly Ile Thr Ser Asp Gln Cys Phe Asp Asn Gly
290          20          25          30
291 Cys Cys Phe Asp Ser Ser Val Thr Gly Val Pro Trp Cys Phe His Pro
292          35          40          45
293 Leu Pro Lys Gln Glu Ser Asp Gln Cys Val Met Glu Val Ser Asp Arg
294          50          55          60
295 Arg Asn Cys Gly Tyr Pro Gly Ile Ser Pro Glu Glu Cys Ala Ser Arg
296 65          70          75          80
297 Lys Cys Cys Phe Ser Asn Phe Ile Phe Glu Val Pro Trp Cys Phe Phe
298          85          90          95
299 Pro Asn Ser Val Glu Asp Cys His Tyr
300          100          105
302 <210> SEQ ID NO: 15
303 <211> LENGTH: 540
304 <212> TYPE: DNA
305 <213> ORGANISM: Homo sapiens
307 <220> FEATURE:
308 <221> NAME/KEY: CDS
309 <222> LOCATION: (41)...(292)
311 <400> SEQUENCE: 15
312 atccctgact cggggtcgcc tttggagcag agaggaggca atg gcc acc atg gag      55
313          Met Ala Thr Met Glu
314          1          5
316 aac aag gtg atc tgc gcc ctg gtc ctg gtg tcc atg ctg gcc ctc ggc      103
317 Asn Lys Val Ile Cys Ala Leu Val Leu Val Ser Met Leu Ala Leu Gly
318          10          15          20
320 acc ctg gcc gag gcc cag aca gag acg tgt aca gtg gcc ccc cgt gaa      151
321 Thr Leu Ala Glu Ala Gln Thr Glu Thr Cys Thr Val Ala Pro Arg Glu

```

VERIFICATION SUMMARY

DATE: 04/17/2001

PATENT APPLICATION: US/09/313,434A

TIME: 14:14:03

Input Set : A:\00786-432001.TXT

Output Set: N:\CRF3\04172001\I313434A.raw